DOCUMENT RESUME

ED 161 937 TM 007 972

AUTHOR McKee, Barbara G.; Hausknecht, Michael A.
TITLE Classroom Assessment Techniques For Hearing Impaired Students: A Literature Review.
PUB DATE [Mar 78]
NOTE 38p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978)

EDRS PRICE MF-$0.83 HC-$2.06 Plus Postage.
DESCRIPTORS *Achievement Tests; *Aurally Handicapped; Higher Education; Item Analysis; Language Handicaps; *Literature Reviews; Response Style (Tests); Secondary Education; Test Bias; *Test Construction; *Testing; *Test Items
IDENTIFIERS *Teacher Developed Tests

ABSTRACT

Literature on the development of classroom achievement tests for high school and college level hearing impaired students is reviewed, with emphasis on achievement tests designed to ascertain whether a particular unit of instruction has been understood as it was presented. The paper reviews: the syntactical structure and vocabulary of test items; item format; modes of delivery of test instructions and items; test-wiseness and propensity for guessing; and non-paper and pencil assessment techniques. The authors contend that research has suffered from an emphasis on the written and spoken language production of hearing impaired students, as opposed to the syntactical structures and vocabulary which they can accurately comprehend. They recommend a number of strategies for testing hearing impaired students: (1) confine test items to a vocabulary level appropriate for students; (2) avoid complex and compound sentences; (3) minimize complex negation formations and items in which the correct answer is a negative answer; (4) be aware that fill-in-the blanks and short answer items are often easier for students to understand than multiple-choice or matching items; (5) include complete written instructions and reinforce them with manual communication; and (6) use visual aids to avoid narrative descriptions. Suggestions for future research are also included. (Author/JAC)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U S DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Barbara McKee

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC) AND USERS OF THE ERIC SYSTEM."

ED161937

CLASSROOM ASSESSMENT TECHNIQUES FOR HEARING IMPAIRED STUDENTS: A LITERATURE REVIEW

Barbara G. McKee
Assistant Professor and Evaluation/Testing Specialist

Michael A. Hausknecht
Evaluation and Testing Assistant

National Technical Institute for the Deaf
Rochester Institute of Technology
Rochester, New York

Presented at the

Annual Meeting of the

American Educational Research Association

Toronto, Ontario

1978

IM007 972

CLASSROOM ASSESSMENT TECHNIQUES FOR HEARING IMPAIRED STUDENTS: A LITERATURE REVIEW

Abstract

As teachers and researchers, few of us can escape the problems of assessing student learning, whether it be a unit of material we have just attempted to teach our students or a question concerning the efficacy of a particular instructional technique we are investigating. Sooner or later we are forced to face the question: Did our students learn what we intended them to learn? This presentation and accompanying paper will report the results of a literature survey with the goal of stimulating discussion about, and raising question concerning, current practices and methodology in the development of classroom achievement tests.

The research reported in this paper was carried out under an agreement with the U.S. Department of Health Education and Welfare.

Introduction

As teachers and researchers, few of us can escape the problem of assessing student learning whether it be a unit of material we have just attempted to teach our students or a question of the efficacy of a particular instructional technique we are investigating. Sooner or later those of us involved in education are forced to face the question: Did our students learn what we intended them to learn? For a variety of reasons, most of us turn to some sort of paper and pencil device to answer this question. Further, as many times as not, we find ourselves developing our own tests because previously developed instruments do not seem to precisely fit the content or student population with which we are concerned.

This paper is directed primarily to those who are interested in developing classroom achievement tests for high school and college age hearing impaired students. The paper reviews the current state of knowledge about item syntax, vocabulary, and format, as well as several other related topics. The emphasis is on classroom tests where the content in question is other than language skills. In other words, the findings presented in this paper might not be employed if the objective of the instrument was the assessment of the students' comprehension of a given syntactical form. However, we are often placed in a position where the content under question is generally unrelated to language skills, e.g. mathematics science, machine use and safety, etc. and we wish to test the students' knowledge of the content presented, not their language skills.

As with any complex question, there are many sides to this coin.

The authors of this paper realize that there are some who would say that similar test materials encountered in real life are not carefully prepared to take into account the particular problems of deaf individuals; and therefore we are perpetrating a disservice on deaf students by attempting to shield them from the realities of modern testing. Examples of frequently encountered tests would be written driving tests and social service exams. However, we can see little credence in this viewpoint. When our goal is not to evaluate students' language skills but to ascertain whether a particular unit of instruction has been understood as it was presented we face a dilemma. We will not know if poor performance on a test indicates a lack of knowledge of the desired facts or a lack of comprehension of the test items; unless we carefully construct our instrument taking into consideration the unique problems of the population we are attempting to teach and evaluate. The problem of preparing the deaf student to deal with the functionings of modern society is a separate issue and should not be allowed to confuse the problems of the measuring of understanding of particular concepts and procedures on the part of deaf students.

The vast majority of research concerned with testing of the deaf has concentrated on the problems of assessing the expressive (speaking and writing) language abilities of deaf students and adults. Recently however, the Office of Demographic Studies at Gallaudet College has extensively examined the results of nationwide standardized achievement tests administered to deaf students in an attempt to evaluate reading comprehension (among other cognitive skills).

The reviewers have attempted to synthesize the information from previous research that would be pertinent to teachers and researchers attempting to construct their own achievement measures.

Although many of the findings reported in this paper are applicable to student surveys, rating scales, and opinion measures; for the most part, we have concentrated on achievement tests i.e., those instruments with a right or wrong answer.

The paper is divided into the following sections:

-Review of the literature relevant to syntactical structure and vocabulary of test items;

-Review of the literature relevant to item format;

-Review of the literature relevant to modes of delivery of test instructions and test items;

-Review of the literature relevant to test-wiseness and propensity for risk taking;

-Review of the literature relevant to non paper and pencil assessment techniques;

-Summary and suggestions for achievement test development;

-Suggested directions for future research.

The information presented under the above headings is meant to be representative of current knowledge on the given topic. The studies cited, however, are not meant to be all inclusive, but representative.

Syntactical Structure and Vocabulary Level

Research in this area suffers (at least from the perspective of this paper) from an emphasis on describing the written and spoken language "production" of hearing impaired students as opposed to a description of the syntactical structures and vocabulary that one would expect could be accurately comprehended by hearing impaired students.

Cooper and Rosenstein (1966) summarized fifty years of research on the language of the hearing impaired by dividing the studies they examined into Free and Controlled methods of data collection.

The free method of data collection characterized those studies that examined students' production of the English language, most often in a written form, although occasionally spontaneous speech was analyzed. These studies were summarized by Cooper and Rosenstein (1966); they found that deaf children wrote significantly shorter sentences, with less diversity of vocabulary, and simpler sentence structure, than hearing children of the same age.

Most traditional studies of the controlled type examined the performance of deaf children on standardized achievement tests, often comparing the results of hearing impaired children to those of hearing children. Cooper and Rosenstein noted that most studies found that the reading comprehension skills of deaf 18-year olds and deaf adults rarely rose above those of the average fourth-grade hearing child.

One of the largest studies to date (DiFrancesca, 1971) has confirmed the findings of Cooper and Rosenstein. Approximately 17,000 deaf students enrolled in primary and secondary programs for the hearing

impaired were given the Stanford Achievement Tests in the Spring of 1971. "Reading Comprehension turns out to be the most difficult academic area of all those covered by the Stanford tests" (p. 9). On the Paragraph Meaning subtest, deaf students progressed from a grade level of 1.6 at six years of age to a grade level of 4.2 at age 18. Only 5.3% of the 18-year-old deaf students reached the 75th percentile for their age (based on hearing norms) on this subtest.

Other recently conducted studies have suggested that the results of the earlier inquiries may have given inflated results. Moores (1967) tested the language skills of deaf students using a Cloze procedure. In the Cloze technique students are presented with passages of prose with every nth word deleted, they are required to supply appropriate words. Moores was able to demonstrate, using this technique, that even the low scores received by deaf children on standardized language and reading tests may be an inflated measure of their understanding of complex syntactical structures.

O'Neill (1973) found deaf students to be inferior in their comprehension of "base structure" rules of grammar and concluded that a reading score achieved by a deaf child does not insure his success in the use of materials appropriate for hearing children at the same reading level.

Conley (1976), using a modified Cloze procedure, found deaf children to have significantly more problems with the use and interpretation of idioms in the English language than did their hearing peers. She noted that idiomatic expressions comprise approximately two-thirds of the English language and that idioms are prevalent in children's

reading material (p. 381), yet even when deaf and hearing children are matched for reading comprehension level, the deaf students score significantly lower in their understanding of idioms.

"The 1960's saw the beginning of a new series of analyses that was more concerned with intra-sentence and intra-clause phenomena and with explicating the rules that children use to produce sentences at various developmental levels." (Quigley et al., 1976, p. 25). A major project in this area was conducted by Quigley and his associates over the past several years. They tested 450 prelingually deaf students from across the nation on most of the major aspects of English syntax using his Test of Syntactic Ability (Experimental Editdon). The particular deviations assessed were developed from earlier studies that analyzed the written language of deaf children. His findings seem important enough to summarize in some detail.

Relativization. To form a relativized structure, one sentence must be reduced and embedded within another. An example is: The man is my friend. The man lives down the street. These two sentences can be combined to The man who is my friend lives down the street (p. 71).

Quigley found deaf students had difficulty comprehending relativized sentences, their difficulty with certain types of relativization persisted through the age of 18. "The difficulty deaf students have in understanding relativized sentences, even at 18 years of age, is one indication of the need for control of syntactic structures in reading and other curriculum materials used with these students" (p. 87).

Conjunction. In English, two or more sentences may be combined into

one compound sentence by conjunctions such as and, but, and or. Mary is going home. I will meet her there. These sentences can be combined to: Mary is going home, and I will meet her there. Or, Ellen woke up late. Ellen missed her train. Ellen woke up late and missed her train (p. 88).

Deaf students were found not to have the extreme difficulty with conjunctions that they had with relativization. Conjoined sentences were easier for the students to understand than conjoined subjects, objects, or verb phrases.

Complementation. A difficult recursive process in English is complementation. Customarily, two simple sentences are joined together by a complementizer so as to form a single complex sentence. For example: John knows that Mary is my sister. Here, two simple sentences are joined by that. (John knows it. Mary is my sister.) In this case the appearance of that is optional (p. 107).

Deaf students were found to have a great deal of difficulty with this type of construction. Their near-chance level of performance indicates that they were unable to discriminate between correct and incorrect forms and therefore were forced to guess which construction was correct (p.114).

Pronominalization. Pronouns are used as a means to reduce redundancy by eliminating features of the noun phrase of which the reader or listener is already aware. There are four main requirements which must be satisfied if a pronoun is to be judged acceptable. They are: case, number, person, and in the case of the third person, gender, as well.

In general, Quigley's results indicated that for deaf students, subject and object case pronouns were easier than possessive and reflexive

pronouns (p. 134). The younger deaf students performed fairly poorly on pronominalization, but at the older ages the rules appeared to be fairly well established (p. 136).

Question Formation. There are two major categories of questions: yes/no questions and who, what, when, where, how, and why questions. Yes/no questions can usually be answered with a simple yes or no. Wh-questions require more than a yes or no response, usually at least a phrase or sentence because the subject and auxiliary are often inverted or the wh-word itself is the subject of the interrogation (p. 137).

The Quigley results indicated that deaf students have difficulty in comprehending questions and in judging the grammaticality of question forms. There was a steady increase in comprehension of both question types with age; however, it was only with yes/no questions that the level of comprehension approached that of the younger hearing children (p. 148).

Negation. The acceptable process of negationrrequires that the negative element be placed in the proper position in the verb phrase.

Quigley's study indicated that deaf students have basically mastered the process of negating sentences by the age of 16 to 18 years. Their performance is quite good, especially when compared to other syntactic structures such as relativization. "The hypothesis that deaf students' acquisistion of some aspects of syntax parallels that of hearing children, but at a greatly reduced rate, is upheld in this instance" (p. 166).

The Verb System. A verb has been defined as a word which expresses an action or a state of being; it provides the core of the sentence. There are two kinds of main verbs: Ordinary verbs which either do or

do not require an object, and linking verbs which link the subject to a modifier in the predicate. In addition to two types of main verbs, there are also auxiliary verbs which are used together with main verbs to indicate tense, voice, and syntactic structures (p. 167).

Quigley found that deaf students have considerable difficulty with the English verb system. Three problem areas are noted: Auxiliary verbs, tense sequencing, and be vs. have. They did not have trouble with verb deletion, an area thought previously to present a good deal of difficulty for the deaf student, but they did have considerable trouble with verb agreement as to both number and tense.

To summarize, Quigley found the most difficult structures for deaf students to accurately comprehend were pronominalization, the verb system, complementation, and relativization. They had the least difficulty with negation, conjunction, and question formation (p. 184). Most importantly, it should be remembered that the gap that exists between the standard English of common reading material and the deaf child's language is very large and represents a significant impediment to their learning even the most basic content from standard texts.

Other studies (Cooper, 1967; Schmitt, 1968; Taylor, 1969; Marshall and Quigley, 1970; Power, 1971; O'Neill, 1973) have used a conceptual base akin to that used by Quigley and have found results similar to those reported by him. However, some researchers have taken a different path. Bochner (1978) has criticized the use of conventional grammatical models of English with their emphasis on normal development in hearing children to characterize deviant language in deaf children and adults.

> In sum, describing the general principles which organize the grammar and underlie the aberrant linguistic behavior in a pre-lingually hearing-impaired population (or other population) entails describing the grammatical systems internalized by members of the group on their own terms and not on the basis of error typologies and/or grammatical and acquisition models of standard English. (Bochner, 1978, p. 5)

Bochner hypothesized that many deaf individuals have internalized grammatical systems which consist of linear linkages of lexical items and do not specify subordination or superordination. Bochner dubbed this hypothesis "The Linear Order Hypothesis," and has completed a study to test his hypothesis. He summarized his findings by saying:

> "...the data indicate that many deaf individuals do not have an adequate knowledge of grammatical relationships which may occur among clauses, consequently, the individuals tend to arrange clauses in a linear-sequential fashion without specifying relationships of subordination"...(p.30).

Bochner qualified this finding by saying his study indicated that approximately 1/3 of his subjects could be characterized by this hypothesis, less than 1/3 could not, and for slightly over 1/3 of the study group, it was impossible to determine whether or not they could be so characterized.

Bochner's summary of the implications of his findings for the development of instructional materials for the deaf (including assessment materials) is summarized in the following citation:

> The major effect the anomolies discussed in this paper have on language use is in one's ability to process (read) unfamiliar information; i.e., information out of the realm of one's own personal experience and knowledge of the world...if one is not familiar with the context of a discourse or its subject matter, he must rely completely on linguistic cues and be able to perceive the logical relationships which obtain among clauses via such cues. In sum, it is likely that deaf individuals who perceive complex sentences in a linear-sequential fashion will have difficulty processing unfamiliar information accurately since logical relationships obtaining among clauses will not be readily accessible to them. (p. 328)

Vocabulary. The studies mentioned thus far have been primarily concerned with sentence structure or, in the case of reading comprehension scores, a combination of sentence structure and vocabulary.

However, several studies have examined the vocabulary level of deaf students apart from the syntax of sentences. Cooper and Rosenstein (1966), in their extensive review, stated that the vocabulary level of deaf 18 year-olds was judged to be about the fourth grade level. DiFrancesca (1972) found similar results in the nation-wide testing program conducted by the Office of Demographic Studies at Gallaudet College.

Simmons (1962) found deaf students to be more redundant in their use of words than hearing students (i.e. using fewer different words in a passage of set length).

Other studies have also found that hearing impaired people are restricted in terms of flexibility and breadth of vocabulary usage (Rosenstein and MacGinities, 1969; Silverman-Dresner and Guilfoyle, 1972).

Walter (1977, 1978) developed a technique to measure the vocabulary comprehension of deaf students based on the idea that their vocabulary knowledge can be directly related to the frequency of occurance in print of any particular word. He used the American Heritage Word Frequency Book which lists, in order of frequency of occurance, over 25,000 words. Walter found that for 14 year old deaf students to read correctly 66 percent of the words given them, the words must fall within the 2,000 most frequently appearing words according to the Heritage List.

Earlier use of the technique with college-age students (NTID students) indicated that in order to maintain a mean of 66 percent correct interpretation, the words must fall within the 10,000 level in frequency.

In summary, this section appears to offer some strong suggestions to the classroom teacher attempting to construct classroom paper and pencil tests of subject matter not related to reading and writing skills:

(1) Confine yourself to a vocabulary level appropriate for your pupils:

Jr. high age - 2,000 on the American Heritage List,
College age - 10,000 on the American Heritage List.

(2) Avoid the following grammatical structures:

Compound and complex sentences in test items and test directions, minimize the number of wh-questions, and complex negation formations.

Item Format

Researchers interested in the problems of assessing the knowledge of deaf students have only recently begun to evidence an interest in the problems concerned with the item format used in tests with deaf students.

Trybus and Buchanan (1973), while examining patterns of achievement in a nationwide administration of the Stanford Achievement Test, noted two special problems in the course of their analysis.

Position of the Correct Response. Traditionally, in multiple choice tests the position of the correct answer is randomized among the alternatives, so that in a given test the correct answer will appear in the various positions an equal number of times. It is expected that the performance of the groups would not be affected by the positioning of the correct response. This, however, was found not to be the case in the analysis completed by Trybus and Buchanan. While the specific position showing the best or poorest performance among hearing impaired students varied from sub-test to sub-test on the Stanford Achievement Test, there seemed to be a definite difference due to positioning. They found that for the Intermediate I battery, the first two positions produced the best performance in 12 out of the 14 sub-tests. In the advanced battery, positions 1 and 4 were best 11 times out of 13. Poorest performance was more evenly distributed in both cases. Trybus and Buchanan stated: "Some experimental study and the evaluation of various kinds of test-taking training will be needed to clarify the issues raised by these findings" (p.18).

Negative Responses. The second special problem noted during the Trybus and Buchanan analysis of the Stanford data is that of the students responses to items in which the correct answer is a negative answer. The form of the negative responses were of the type:

"N" Neither alternative is correct

"NP" No punctuation needed

"NG" Correct answer is not given

Hearing impai[illegible]tudents, according to the Trybus and Buchanan analysis, tended to do substantially less well than their hearing peers on items where the correct answer is one of the "negative" responses given above. "Full explication of the roots of this phenomenon must await experimental studies focused on this issue" (p. 18).

Matching vs. Multiple Choice Test Format. Walter (1976) administered two forms of a vocabulary test to two groups of entering NTID students during the summer of 1976. All incoming students were randomly assigned to one of the two groups. It was demonstrated that the two groups did not differ significantly on measures of vocabulary, reading, and written language competence. One form of the test used an 8 sentence, 10 word matching format. The other form employed the same words and sentences but in a multiple choice format with five options.

Student responses on the two formats were compared. Walter found that the mean scores for the matching format and the multiple choice format did not differ significantly. Hence the use of either format did not appear to influence the students' total scores. Walter decided to continue exploring only the multiple choice form of the test

because this form was easier to administer and score and took students less time to complete.

Recall vs. Recognition Items. A study by Osguthorpe and Long (undated) measured the effects of reviewing class notes on four learning outcomes: Recognition, recall, concept acquisition, and problem solving, for a group of 53 deaf and 33 hearing college age students. An unexpected result that surfaced in the study was the performance of the deaf students on the recall and recognition items. The recall test consisted of 12 definitions. Students were required to recall the term for each definition and write the term beside its definition. The recognition portion consisted of the same 12 items found in the recall portion except that the definitions had been randomly reordered and the correct terms were given. The students were required to match the term with its definition. Osguthorpe and Long summarized their findings on this point thusly:

> The individual means show that hearing students nearly doubled their performance from recall (m=5.8) to recognition (m=10.1) while deaf students increased only slightly, recall (m=4.48), recognition (m=6.05). Further analysis showed that approximately half (47%) of the delayed review deaf students' recognition scores were equal to or actually lower than their recall scores...In other words having the response set available was highly facilitative for hearing students, but not for deaf students. Many deaf students were actually confused by having the correct answer on the test form. This finding has been replicated in a subsequent experiment (Ellsworth, 1977) (p. 11, 12).

Item Bias. In a recent article, Rudner (1978) applied a technique suggested by Angoff (1972) to determine which items contained in 12 subtests of the Intermediate I battery of the Stanford Achievement Tests were biased for or against deaf students. His sample consisted of 618

normally hearing and 2,821 hearing impaired examinees. The hearing-impaired subjects were some of those tested in the 1971 Annual Survey of Hearing Impaired Children and Youth conducted by the Office of Demographic Studies at Gallaudet College.

Rudner identified 48 items (out of 419) that were biased in favor of one of the two populations. Of these 48 biased items, 22 appeared to be biased in favor of hearing impaired examinees and 26 in favor of hearing examinees. As Rudner pointed out, since the number of biased items is low and since the biased items are fairly uniformly distributed between population groups, one may suspect that there is little bias in total scores. However, Rudner pointed out two reasons this may not be the case:

1) First, the limitations of the procedure may lead to spurious conclusions. Indices of item difficulty i.e. p. values, are obtained for each item under investigation for each group. These p. values are then converted to Z-scores and pairs of Z-scores are plotted on a 2-axis graph. Items deviating significantly from the main axis of the plot are considered to be biased. Item difficulty may not be the only characteristic that should be considered in the determination of item bias.

2) In this study, the hearing impaired subjects were, on the average, five or six years older than the hearing subjects. Students were assigned to a particular battery based on reading levels. Because of their higher mean age (and hence greater repertoire of experience), the hearing impaired students might do well even on items biased against them. Further (said Rudner) one would anticipate several items being identified as biased in favor of the hearing impaired because of this age and experience difference, rather than because of item characteristics.

Rudner classified the items biased against the hearing-impaired subjects by the linguistic structure of the items and found the results to conform with earler work of his (and also with other studies reviewed in earlier sections of this paper).

He found that the following linguistic structures appear to be misleading for hearing-impaired students:

1. Conditionals (if, when);
2. Comparatives (greater than, the most);
3. Negatives (not, without, answer not given);
4. Inferentials (should, could, because, since);
5. Low information pronouns (if, something);
6. Lengthy passages

Methods of Delivery

The strength and weakness of the various methods of presenting instructions for a test, and in some cases the test items themselves, has not been extensively examined in the literature dealing with the assessment of the hearing impaired. Most studies in this area have concentrated on the examination of different modes of presenting learning materials and usually the criterion measure is some sort of paper and pencil test. Unfortunately, the psychometric properties of the criterion test are not always examined.

Panko (1975) compared four types of mediated verbal presentation formats: 1) an unaltered videotape designed for hearing students, 2) the same videotape with captions, 3) the videotape with live manual translations, and 4) a print version of the visuals replacing the charts and graphs of the videotape. The criterion measure used was a 26 item multiple choice test. Panko found that students scored the highest when given the printed text, but the difference was significant only when compared with the unaltered videotape and was not significant when compared with either the captioned or interpreted versions. Surprisingly, there was no significant difference in achievement noted between students who viewed the captioned version of the videotape and students who viewed the unaltered videotape. This could have been due to the graphic nature of the material presented.

In his literature review on the use of captioning, Panko made the following comment:

> Little research has been conducted testing the effects of captions on post secondary deaf students. Almost no research has been conducted comparing the effects of various presentation modes when the content is judged to be on the college level. (p. 38)

Caccamise et al. (undated) compared the achievement of three groups of deaf students on a true-false test of factual information after they were presented a lecture containing the information in three different presentation modes; live lecture using simultaneous communication, direct transmission television, and direct transmission rear screen projection. The group viewing the live lecture achieved significantly higher scores on the criterion measure than did the groups viewing the visual media transmissions.

Stuckless (undated) in a paper discussing the application of technology to deaf education reviewed several studies that investigated different visual modes of presenting verbal information to deaf persons (Gates, 1971; Propp, 1972; Norwood, 1976; Stuckless and Pollard, 1977). Stuckless concluded that "The results of both basic and applied research.. ..furnish conditional support for the use of print for receptive communication by deaf persons" (p. 11).

Bradinger (1970) gave three groups of randomly assigned deaf students instructions for a paper and pencil standardized test by three different methods: 1) Specially written programmed instruction booklets, 2) Manual communication, 3) Oral directions plus the written directions included on the test. She found no significant difference in test performance between the three groups.

Ross (1967) studied the effects of presenting the questions on the Wechlser Adult Intelligence Scale by both the simultaneous communication method and the oral-written method. He found that his subjects (48 manual students, ages 16-21) performed significantly better when they were presented the questions in the simultaneous communication mode.

Higgins (1973) compared performance on a multiple choice test of three groups of Gallaudet students after they had viewed two brief videotapes, one group employing total fingerspelling (The Rochester Method), one group employing American Sign Language (ASL), and the last employing Siglish (a form of manually coded English). The Siglish method was found to lead to significantly better performance on the criterion measure than either the ASL or fingerspelling methods of communication. It was pointed out that the criterion measure was written in standard English (hence, Siglish more closely approximated the criterion measure), and Gallaudet students are generally accustomed to some form of manually coded English through their classroom work.

Hoemann (1972) wanted to test the accuracy of interpreted communication in a group situation. At a workshop for interpreters for the deaf, a simple ten item test was both read aloud and signed to the participants. All the participants were adults and the deaf participants were invited consultants, respected by both the deaf community and the Interpreter Organization (an average or above average intelligence for these subjects can be assumed). The questions were such that any adult with normal intelligence would have little difficulty in answering them correctly, e.g., In what month does Christmas fall? (There was one nontrivial question included in the group of ten to lend credulity to the statement that this was a test). The mean score of the 16 hearing subjects was 9.1 (n=16), the score of the hearing impaired participants was 5.6 (n=9). Hoemann's study points out the problems of receiving manual communication in a group setting even when the receivers are motivated adults.

Hitti, Riffer and Stuckless (1971) administered the Raven's Progressive

Matrices general aptitude test to two groups of randomly selected incoming NTID students. The first group completed the test under conventional group procedures, the second group took the test under computer-managed conditions where directions were given in graphic form on a cathode ray tube. The students were seated at individual consoles. They still had to read the test items from the test booklet, but they recorded their answers on the typewriter keyboard of the console. Using the Computer Monitored Instruction (CMI) method, the student had only one chance to change his answer (immediately after he pressed a key for the answer, a message on the cathode ray tube asked him if he wished to change his answer). Under the conventional method, the students could change responses on their answer sheets for the duration of the test. Twenty-one days later, the students were retested with the groups reversed in presentation mode.

Hitti et al., found no significant difference between the performance of the students under the two modes of test administration. The authors summarized their findings with the comment, "This would suggest that the introduction of the computer to the actual test administration neither interferes with, nor enhances, a student's test performance" (p. 16).

Test Wiseness and Risk Taking

Two aspects of testing that are receiving quite a bit of attention in the literature concerned with factors impacting on test scores (for hearing population) are those of test-wiseness and risk taking behaviors. This area has received almost no attention in the literature concerned with assessment of deaf students. However, one study (Koehler, Kelly and Gonter, 1977) examined these phenomena in a hearing impaired population. Test-wiseness (TW) is defined in the Koehler study as "a subject's capacity to utilize the characteristics and formats of the test and/or test taking situation to receive a high score."

Risk taking (RT), on the other hand, is the tendency to respond to a test item when the correct answer is not known. Risk taking, in testing, is generally limited to those situations in which the student has been informed he or she will be penalized for guessing. According to Koehler, et al. there are data that indicates low risk takers tend to be penalized on their test scores compared to their more courageous associates.

Koehler and his associates evaluated the test-wiseness ability of deaf students using items in which the correct answer could be obtained only by eliminating the incorrect answers or by in some way determining the relationship of the stem to its responses. The questions were essentially content free, in that there were no correct answers. Sixteen such items were embedded in 28 legitimate multiple-choice items.

Risk taking behavior was measured by 5 word-pair items in which the examinee was asked to indicate whether a nonsense word had the same

or the opposite meaning as a real word. These five RT items were embedded in 5 legitimate items.

Koehler found that deaf students were significantly lower in TW ability than their hearing counter-parts, and that the degree or amount of TW did not increase with grade level or reading level as did that of hearing students, however, they did consistently receive scores above chance, hence some use of test-wiseness was being applied.

The propensity for risktaking showed far less difference between deaf and hearing subjects. Although 9th grade hearing subjects were substantially higher risk takers, the difference became less in the 10th and 11th grades and disappeared in the 12th grade.

Koehler and his associates pointed out that the lesser proficiency in language of the deaf students would affect both their acquisition and demonstration of test-wiseness. These reviewers examined the items for the 5th grade test-wiseness section and felt that although the stems were indeed content free, the vocabulary level of the stem and options, and at times, the content of the options, would have prevented many deaf students from applying test-wiseness skills whether or not they possessed them.

Non Paper and Pencil Assessment Techniques

One of the most underexamined (or at least unpublished) topics in the general area of assessment of classroom learning of deaf students is that of non-paper and pencil assessment devices. Certainly teachers have used various types of observation techniques to determine when their students have mastered certain skills, but seldom have such techniques been systematically examined and even more seldom, are traditional achievement tests compared to other methods of assessing subject mastery.

Vernon and Brown (1964) strongly recommended that all I.Q. tests administered to deaf children be non-verbal performance instruments. They have backed up their recommendation with examples of children severly misdiagnosed on the basis of verbal I.Q. tests. A later review of 50 years of research on intelligence testing of deaf persons (Vernon, 1968) reinforces the fact that when deaf children are tested by appropriate methods and have no additional handicapping condition "The deaf and hard-of-hearing function at approximately the same I.Q. level on performance intelligence tests as do the hearing."

A project initiated at Gallaudet College (Roy, 1961) attempted to construct and norm a Picture Interest Inventory (PII) for use in vocational counseling situations. The inventory consisted of twenty-seven picture triads, each picture in the triad was judged to represent a different occupation or activity.

Several efforts were made to validate the Picture Interest Inventory. One such effort compared subjects' scores on the PII to the subjects'

scores on the Kuder Preference Record. The correlation of the 9 scales of the PII which corresponded to 9 scales of the Kuder Record varied from .26 to .62. Similar efforts to compare pictorial and verbal achievement tests (other than I.Q. tests) were not found.

The NTID Visual Communication department has developed an extensive and relatively objective student assessment plan for their Printing Production Program that does not rely on paper and pencil assessment instruments. The faculty developed specific skill objectives and required that the students demonstrate mastery of the skills in a laboratory situation. Such a technique is similar to the competency based curriculum approach now receiving so much attention in higher education. One aspect of occupational education that might be considered different from liberal arts education is the need for the student to reach a performance level which approximates 100%. Hence the students are simply rated on whether or not they can successfully complete the task, the only variable remaining is the time needed to reach that level of performance. Such time data is now routinely collected in the Printing Program.

Preliminary results of the time data indicate a possible bi-modal distribution "There is an implication that students who do not complete a skill or skills within the normal time distribution will never complete a course unless comething is done to change the students' learning success." (p. 30) Most certainly this is an area that deserves additional examination.

Summary and Suggestions for Test Developers

This paper reviewed the literature relevant to teachers and researchers concerned with developing classroom assessment instruments for hearing impaired students.

The first section condensed the voluminous amount of literature concerned with the syntactical and vocabulary comprehension skills of deaf students. Suggestions to achievement test developers arising from this section were:

1) Confine yourself to a vocabulary level appropriate for your pupils; for high school and college age students this would be somewhere between the 5,000 and 10,000 word level on the American Heritage Word List.
2) Avoid, whenever possible, compound and complex sentences in test items and test directions. Minimize the number of wh-questions and complex negation formations.
3) There is a possibility many of your students will be interpreting phrases in a linear fashion, take this into account when developing test items.

The second section reviewed the available literature on the item formats that appeared to present difficulty to many hearing impaired students. Suggestions to test developers arising out of this section were:

4) Minimize items in which the correct answer is a negative answer.
5) Be aware that a multiple choice or matching format (recognition) item may not be any easier for your students than a fill-in-the blank or short answer item requiring a one word response.

6) Caution your students to read the entire item (all options) carefully before responding.

The next section dealt with methods of presenting test directions' and test items. On this topic the research is unclear, however, a few suggestions seem to emerge:

7) Always include complete written instructions, in addition, use manual communication to reinforce and elaborate on the written instructions.

8) If possible, use visual aids to avoid lengthy narrative descriptions in both test directions and test items.

The research conducted thus far concerning the test wiseness and risk taking characteristics of hearing impaired students is so minimal that no conclusions or suggestions can be offered.

The final topic covered in this paper was non-traditional methods of assessing the learning of deaf students. Again, so little has been documented that no conclusions can be drawn or recommendations offered. Clearly, this is one area that desparately deserves further study.

This paper did not review the traditional literature on constructing classroom achievement tests. It goes without saying that a carefully drawn test specification grid should be developed that lists the content areas to be covered and the types of questions to be included (e.g. recall items, recognition items, concept acquisition, problem solving, etc.). The reviewers have included several sources in the bibliography that are of a general nature for those readers, unaware of this literature, who wish to pursue the topic further. These general sources are indicated in the bibliography by an asterix preceding the author's name.

In summary, the authors would like to make a recommendation to test developers not mentioned in the literature. Conduct a pilot test for any classroom assessment technique that will be used repeatedly, or that will carry above average importance. The subjects for such a pilot test could be students who have previously completed the course or who covered elsewhere the content assessed by the test. If necessary, the first use of the test in the classroom could be considered the pilot, with little weight given to the test when determining the student's grade. After the pilot test, not only should the students' responses be carefully examined, but the test or assessment technique should be discussed with the students as well. Issues to be considered might be: How did the students interpret a given question? Why did the students choose a particular response, or give a particular answer?, etc. These and similar questions should be explored with the group of students after their completing the test.

Suggestions for Future Research

Somewhat distressingly, the research review presented in this paper raised more questions in the minds of the reviewers than it answered. In general, any additional studies in the area covered by this review should be of great significance to those concerned with the educational assessment of hearing impaired individuals. The following comments occurred to us:

1) The reviewers would like to strongly encourage Bochner and others to continue the investigation of his Linear Hypothesis Theory.
2) More study is needed on the vocabulary comprehension skills of hearing-impaired students. For example, is the American Heritage Word List the most feasible instrument to use, or, should a modification of such a list be developed for the use of those designing instructional materials and assessment techniques for use with the hearing impaired. The question of deaf students' comprehension of second and third meanings of words needs further exploration.
3) More studies need to be conducted examining the effect of the positioning of the correct response in multiple choice questions.
4) Further work should be undertaken examining the item formats in objective tests.
5) An insufficiently unexplored area is the test-wiseness ability of deaf students.

6) The problem of congruence between the method of delivery of information and the method of its assessment has not been examined. Several studies have examined different modes of presenting information, but the criterion measure was almost always a paper and pencil test in standard English.

7) More methods employing the demonstration of desired skills need to be developed and examined. For example, the NTID Printing Program assumes that a student either can or can not perform a given task. Is this truly a dichotomous variable?

REFERENCES

Adler, E. P., Research Trends in Deafness: State of the Art. Washington, D.C: Department of Health, Education, and Welfare, June, 1970.

Allen, D. V., Color-Word Interference in Deaf and Normal Children. Paper presented to the Midwestern Psychological Association, Detroit, 1971.

*Anderson, R. C., How to Construct Achievement Tests to Assess Comprehension. Review of Educational Research, 1972, 42, (2), 145-170.

Birch, J. R., An Eleven Year Study of Predicting School Achievement in Young Deaf Children. American Annals of the Deaf, 1963, 108, (2), 236-240.

*Bloom, B.S., Hastings, J. T., & Madaus, G. F., Handbook on Formative and Summative Evaluation of Student Learning. New York: McGraw-Hill, 1971.

Bochner, J., Error, Anomaly, and Variation in the English of Deaf Individuals. Unpublished Manuscript, National Technical Institute for the Deaf, Rochester.

Bolton, B., A Note on the Prediction Ceiling. Journal of Rehabilitation of the Deaf, 1972, 5 (3) 24-28.

*Bormuth, J. R., On the Theory of Achievement Test Items. Chicago: University of Chicago Press, 1970.

Brandinger, A., Programmed Instruction, Sign Language, and Formal Directions to a Reading Test for Deaf High School Students. Doctoral Dissertation, Rutgers University, 1970.

Brill, R. G., The Superior I.Q.'s of Deaf Children of Deaf Parents. Journal of Rehabilitation of the Deaf, 1970, 4 (2), 45-53.

Caccamise, F., Blasdell, R., & Meath-Lang, B., Hearing Impaired Persons Simultaneous Reception of Information Under Live and Two Visual Motion Media Conditions. Unpublished, Undated Manuscript, National Technical Institute for the Deaf, Rochester.

Conley, J. E., The Role of Idiomatic Expressions in the Reading of Deaf Children. American Annals of the Deaf, 1976, August, 381-385.

Cooper, R., The Ability of Deaf and Hearing Children to Apply Morphological Rules. Journal of Speech and Hearing Research, 1967, 10, 77-86.

Cooper, R. L., & Rosenstein, J., Language Acquisition of Deaf Children. Volta Review, 1966, 68, (1), 58-67.

Costello, E. R., & Root, A. A., Appraising Certain Linguistic Structures in the Receptive - Sign Language Competence of Deaf Children. Report of the Procedings of the 47th Meeting of the Convention of American Instructors of the Deaf. U. S. Government Printing Office, Washington, 1976, 123-128.

Culhane, B. R., The Effects of Written English and Ameslan Adjuncts to Pictures on a Concept Identification Task Among Prelingually and Post Lingually Profoundly Deaf Postsecondary Students. Unpublished Manuscript, National Technical Institute for the Deaf, Rochester, 1975.

DiFrancesca, S., Academic Achievement Test Results of a National Testing Program for Hearing Impaired Students - United States: Spring, 1971. Annual Survey of Hearing Impaired Children and Youth, D, (9), 1-41.

*Educational Testing Service, Making the Classroom Test: A Guide for Teachers. Princeton, N. J., Educational Testing Service, 1973.

Gentile, A., & DiFrancesca, S., Item Analysis of Academic Achievement Tests Hearing Impaired Students United States: Spring, 1969, Annual Survey of Hearing Impaired Children and Youth, D, (2), 1-34.

*Gronlund, N. E., Preparing Criterion - Referenced Tests for Classroom Instruction. New York: MacMillan, 1973.

Higgins, E., An Analysis of the Comprehensibility of Three Communication Methods Used With Hearing Impaired Students. American Annals of the Deaf, 1973, (February), 46-49.

Hitti, F. J., Riffer, R. L., & Stuckless, E. R., Computer-Managed Testing: A Feasibility Study with Deaf Students: Unpublished Manuscript, National Technical Institute for the Deaf, Rochester, 1971.

Hoemann, H. W., Communication Accuracy in a Sign-Language Interpretation of a Group Test. Journal of Rehabilitation of the Deaf, 1972, 5 (3), 40-43.

Isaacs, M., & Hughes, F., PSI as a Supplement to College Classroom Lecture Instruction. Journal of Rehabilitation of the Deaf, 1975, 9 (2), 44-54.

Jensema, C., A Statistical Investigation of The 16PF Form E as Applied to Hearing Impaired College Students. Journal of Rehabilitation of the Deaf, 1975, 9 (1), 21-29.

Kazmierski, P. R., The Effects of the Cloze Procedure Upon the Literal Understanding of Text Materials by Postsecondary Deaf Students. Unpublished Manuscript, National Technical Institute for the Deaf, Rochester, 1973.

Koehler, R. A., Kelly, R. R., & Gonter, M. A., Test-Wiseness and Risk Taking of Deaf and Normal Hearing Subjects. Paper presented at AERA Convention, New York, 1977.

Lane, H. S., & Baker, D., Reading Achievement of The Deaf: Another Look. The Volta Review, 1974, 76 (8), 489-499.

Long, G. L., Conklin, D., & Garrison, W., The Development of a Cognitive Process Based Learning Strategies Questionnaire. Unpublished manuscript, National Technical Institute for the Deaf, Rochester, 1978.

MacGinitie, W. H., Ability of Deaf Children to Use Different Word Classes. Journal of Speech and Hearing Research, 1964, 7, 141-150.

Maher, H., A Comparison of Individualized Independent Instruction and the Traditional Lecture Format as it Effects Learning Experiences of NTID Deaf Students in Biology. Unpublished, Undated Manuscript, National Technical Institute for the Deaf, Rochester,

Marshall, W., & Quigley, S. P., Quantitative and Qualitative Analysis of Syntactic Structure in the Written Language of Deaf Students. Urbana, Ill.: Institute for Research on Exceptional Children, 1970.

Office of Demographic Studies, Gallaudet College. Further Studies in Achievement Testing Hearing Impaired Students - United States: Spring, 1971. Annual Survey of Hearing Impaired Children and Youth, D (13).

O'Neill, M. A., The Receptive Language Competence of Deaf Children in the Use of the Base Structure Rules of Transformational-Generative Grammar. Unpublished Doctoral Dissertation, University of Pittsburgh, 1973.

Osguthorpe, R. T., Long, G. L., & Ellsworth, R. G., The Effects of Reviewing Class Notes for Deaf and Hearing Students. Unpublished, Undated Manuscript, National Technical Institute for the Deaf, Rochester,

Panko, J., The Learning Effects of Visually Mediated Verbal Presentation Formats on Deaf Postsecondary Students. Doctoral Dissertation, University of Pittsburgh, 1975.

Panko, J., The Immediate Learning Effects of Four Visually Mediated Presentation Formats. Unpublished, Undated Manuscript, National Technical Institute for the Deaf, Rochester.

Pollack, C., The Effects of the Oral and Simultaneous methods of Communication on Test Performance Involving Comprehension of Concrete and Abstract Concepts Among Deaf High School Students. Doctoral Dissertation, New York University, 1966.

*Popham, W. J., Educational Evaluation. Englewood Cliffs, N. J.: Prentice-Hall, 1975.

Power, D. J., Deaf Children's Acquisition of The Passive Voice. Unpublished Doctoral Dissertation, University of Illinois, 1971.

Riffer, R., The Validity of Selected Paper and Pencil Instruments for Measuring NTID Students' Vocational Interests. Unpublished Manuscript, National Institute for the Deaf, Rochester, 1971.

Rosen, A., Limitations of Personality Inventories for Assessment of Deaf Children and Adults as Illustrated by Research with the Minnesota Multiphasic Personality Inventory. Journal of Rehabilitation of the Deaf, 1967, 1 (2), 47-52.

Ross, D. R., Test Performance of Deaf Adults Under Two Modes of Test Administration. Doctoral Dissertation, University of Arizona, 1967.

Roy, H. L., Klein, F., & Bornstein, H., Report of Construction of a Picture Interest Inventory for the Deaf, 1959-1961. Unpublished, Undated Manuscript, Gallaudet College, Washington, D. C.

Rudner, L. M., Using Standard Tests with the Hearing Impaired: The Problem of Item Bias, The Volta Review, 1978, 80 (1), 31-40.

Schmitt, P., Deaf Children's Comprehension and Production of Sentence Transformations and Verb Tenses. Unpublished Doctoral Dissertation, University of Illinois, 1968.

Stuckless, E. R., & Enders, M., Three Studies of the Structural Meaning of English for Postsecondary Deaf Students. Unpublished Manuscript, National Technical Institute for the Deaf, Rochester, 1972.

Stuckless, E. R., Technology and Verbal Processing. Unpublished, Undated Manuscript, National Technical Institute for the Deaf, Rochester.

Taylor, L., A Language Analysis of the Writing of Deaf Children. Unpublished Doctoral Dissertation. State University of Florida, 1969.

Trybus, R. J., & Buchanan, C., Patterns of Achievement Test Performance. Studies in Achievement Testing, Hearing Impaired Students United States: Spring, 1971. Annual Survey of Hearing Impaired Children and Youth, D (11), 3-44.

Trybus, R. J., Personality Assessment of Entering Hearing-Impaired College Students Using the 16PF, Form E. Journal of Rehabilitation of the Deaf, 1973, 6 (3), 34-40.

Vernon, M., & Brown, D. W., A Guide to Psychological Tests and Testing Procedures in the Evaluation of Deaf and Hard-of-Hearing Children. Journal of Speech and Hearing Disorders, 1964, 29 (4), 414-423.

Vernon, M., Fifty Years of Research on the Intelligence of Deaf and Hard-of-Hearing Children: A Review of Literature and Discussion of Implications. Journal of Rehabilitation of the Deaf, 1968, 2 (1), 1-12.

Visual Communications Department. A Plan of Work to Implement Individualization of Instruction in Printing Production. Unpublished Manuscript, National Technical Institute for the Deaf, Rochester, 1976.

Walter, G., A Test to Measure Vocabulary Knowledge of Deaf Students. Unpublished Manuscript, National Technical Institute for the Deaf, Rochester, 1976.

Walter, G., A Technique for Measuring Vocabulary Knowledge. Unpublished Manuscript, National Technical Institute for the Deaf, Rochester, 1977.

Walter, G., Lexical Abilities of Hearing and Hearing-Impaired Children. Unpublished Manuscript, National Technical Institute for the Deaf, Rochester, 1978.

*Wesman, A. G., Writing the Test Item. Educational Measurement (2nd Edition). Washington, D. C.: American Council on Education, 1971, Chapter 4.

Wilbur, R. B., An Explanantion of Deaf Children's Difficulty with Certain Syntactic Structures of English. The Volta Review, 1977, 79 (2), 85-92.

*Wood, D. A., Test Construction: Development and Interpretation of Achievement Tests. Columbus, Ohio: C.E. Merrill, 1960.